



**Calhoun: The NPS Institutional Archive**  
**DSpace Repository**

---

Theses and Dissertations

1. Thesis and Dissertation Collection, all items

---

1969

An investigation of the effects of computer graphics on executive decision making in an inventory control environment.

Prokop, Jan Stuart

University of North Carolina at Chapel Hill

---

<http://hdl.handle.net/10945/12426>

---

*Downloaded from NPS Archive: Calhoun*



Calhoun is the Naval Postgraduate School's public access digital repository for research materials and institutional publications created by the NPS community. Calhoun is named for Professor of Mathematics Guy K. Calhoun, NPS's first appointed -- and published -- scholarly author.

**Dudley Knox Library / Naval Postgraduate School**  
**411 Dyer Road / 1 University Circle**  
**Monterey, California USA 93943**

<http://www.nps.edu/library>

NPS ARCHIVE  
1969  
PROKOP, J.

AN INVESTIGATION OF THE EFFECTS OF COMPUTER  
GRAPHICS ON EXECUTIVE DECISION MAKING  
IN AN INVENTORY CONTROL ENVIRONMENT

Jan Stuart Prokop  
//

University of North Carolina  
at Chapel Hill



Department of Computer and  
Information Science

Thesis  
p9444



AN INVESTIGATION OF THE EFFECTS OF COMPUTER  
GRAPHICS ON EXECUTIVE DECISION MAKING  
IN AN INVENTORY CONTROL ENVIRONMENT

by

Jan Stuart Prokop

//

A thesis submitted to the faculty of  
the University of North Carolina at  
Chapel Hill in partial fulfillment of  
the requirements for the degree of  
Doctor of Philosophy in the Department  
of Computer and Information Science

Chapel Hill

1969

Approved by:

NPS ARCHIVE

1969

PROKOP, J.

~~7/28/68~~  
~~P/9444~~

### ACKNOWLEDGEMENTS

This research was made possible by the Ph.D. Studies Program of the Department of the Navy.

The 2250 display unit used was made available under a research contract between the International Business Machines Corporation, Federal Systems Division and the University of North Carolina at Chapel Hill.

The author is deeply indebted to Dr. Frederick P. Brooks, Jr. for his interest, insight, encouragement and generous participation in this experiment.



## PREFACE

The central question of this dissertation is whether an objective ranking of utility to the decision maker can be assigned to the form of the presentation to him. In this instance we are interested in whether an executive decision can be reached earlier, faster or more consistently with a computer-driven display device than with the more customary printed material.

An experimental group of eighteen management-level subjects with extensive experience in inventory control was assembled for a two-week short course in advanced inventory management techniques. During the short course, the subjects were exposed to simulated results from computer application of certain inventory control policies to a hypothetical inventory system handling  $n$  items. This system is faced with a certain randomly derived set of orders, price changes, replenishments of stock and other transactions. The results of the simulation were presented on both printer paper and on a cathode ray tube display device. For each method of presentation, results were represented in both tabular and graphical form.

When a substantial part of the course had transpired, the subjects were asked to evaluate the results of simulating two inventory systems, using printer output for





one evaluation and the cathode ray tube display device for the other evaluation. By means of a Latin square design and rank order statistics these evaluations were inspected to determine if experienced decision-makers using a display device could reach a decision that was earlier, faster or more consistent than a decision reached by means of printer output.

The results indicated, with very high statistical significance, that a decision could be made faster with a display device, and with high significance that a decision could be made on the basis of less information by means of the display device. Other results pointed to decisions that were more consistent for the individual and which tended more to agree with the rest of the group when display device techniques were used for the evaluation of the inventory systems.



## CONTENTS

I. INTRODUCTION . . . . .	1
II. THE SIMULATION MODEL . . . . .	7
III. THE PRESENTATION TO THE DECISION MAKER . . . . .	22
IV. THE SUBJECTS OF THE EXPERIMENT . . . . .	37
V. THE DESIGN OF THE EXPERIMENT . . . . .	45
VI. RESULTS . . . . .	54
VII. CONCLUSIONS AND REMARKS . . . . .	83
SELECTED BIBLIOGRAPHY . . . . .	89



## LIST OF TABLES

Table	Page
1. a. Ordinal Value of Decision Month . . . . .	57
b. Transformed Ordinal Value of Decision Month . . . . .	58
2. Analysis of Transformed Decision Month . . . .	59
3. Elapsed Time in Minutes to Decision Month . . .	61
4. Analysis of Decision Time . . . . .	62
5. Time in Minutes to Complete the Problem After the Decision Month . . . . .	65
6. Analysis of Time to Complete Problem After Decision Month . . . . .	66
7. a. Spearman Rank Correlation Coefficient, $r_s$ .	69
b. Transformed Spearman Rank Correlation Coefficient, $r_s$ . . . . .	70
8. Analysis of Transformed Spearman Rank Correlation Coefficient, $r_s$ . . . . .	71
9. Data and Results from Kendall's W Statistic Computations for Upper Left Quadrant of Latin Square; Group I, Order 1, Display Unit . . . . .	73
10. Data and Results from Kendall's W Statistic Computations for Upper Right Quadrant of Latin Square; Group I, Order 2, Printer Output . . . . .	74
11. Data and Results from Kendall's W Statistic Computations for Lower Left Quadrant of Latin Square; Group II, Order 1, Printer Output . . . . .	75



12.	Data and Results from Kendall's W Statistic Computations for Lower Right Quadrant of Latin Square; Group II, Order 2, Display Unit . . . . .	76
13.	Data and Results from Kendall's W Statistic Computations for Upper Left Quadrant of Latin Square at Twenty-Fourth Month; Group I, Order 1, Display Unit . . . . .	78
14.	Data and Results from Kendall's W Statistic Computations for Upper Right Quadrant of Latin Square at Twenty-Fourth Month; Group I, Order 2, Printer Output . . . . .	79
15.	Data and Results from Kendall's W Statistic Computations for Lower Left Quadrant of Latin Square at Twenty-Fourth Month; Group II, Order 1, Printer Output . . . . .	80
16.	Data and Results from Kendall's W Statistic Computations for Lower Right Quadrant of Latin Square at Twenty-Fourth Month; Group II, Order 2, Display Unit . . . . .	81





## LIST OF FIGURES

Figure	Page
1. Three-Dimensional Representation of Figure 2 . . . . .	23
2. Detailed Listing of Simulation Results by Policy Within Month . . . . .	24
3. Sorted Listing of Simulation Results by Policies Within Month . . . . .	25
4. Bar Graph of Simulation Results by Month Across All Policies . . . . .	26
5. IBM 2250 Cathode Ray Tube Display Device . . .	28
6. Display of Figure 2 . . . . .	29
7. Display of Figure 3 . . . . .	30
8. Display of Figure 4 . . . . .	31
9. Quadrant Graph of Simulation Results by Month Across All Policies . . . . .	33
10. Programmed Function Keyboard . . . . .	34
11. Data Collection Format . . . . .	42
12. Latin Square Design . . . . .	46
13. Analysis of the 2 x 2 Latin Square Design . .	48



## I. INTRODUCTION

The central question of this dissertation is whether an objective ranking of utility to the decision maker can be assigned to the form of the presentation to him. In this instance we are comparing the utility of a cathode ray tube display to printed material, and are interested in whether an executive decision can be reached earlier, faster, or more consistently with the display device than with the more customary printed material.

In order to provide a vehicle for the experiment, and to establish an environment for a discussion of the results, a simulator was written for computer application of certain inventory control policies to a hypothetical inventory system handling n line items. This system is faced with a certain randomly derived set of orders, price changes, replenishments of stock, and other transactions. This simulator is governed by a number of pseudo-random variate generators and by program constants. These parameters are set by the user. An important set of these parameters establishes the management policy for inventory control. A policy is formalized by the selection of a combination of three policy components from among:

1. Two service rules to be used in variable safety level computation.



2. Two smoothing constants to be used in demand forecasting by exponential smoothing.

3. Three purchase quantity computation rules.

A Cartesian product of these three sets gives a total of twelve inventory control policies to be investigated. The simulator exercises each of the twelve policies for a period of twenty-four months of generated transactions. At the end of each of the simulation months, relevant statistics are gathered about the system performance under each policy set. In order to insure uniform interpretation of the statistics, each policy set is exercised under identical initial conditions.

Due to the large number of parameters which are available to the user, the simulation program provides a versatile vehicle for this investigation. The statistics which the program develops are quite extensive, and only part of these data are used in the investigation. Although the simulation program and its results are of interest in their own right, the present investigation uses the simulator as a vehicle rather than the object of investigation.

Based on the data presented to him by the simulation program, the decision maker is expected to make a series of decisions. At the end of each simulation month, he ranks the policies in the order of acceptability to him. Each decision maker thus produces a set of twenty-four ordered lists of twelve policies each, from which conclusions will be drawn concerning the form of presentation of the data.



The experiment was conducted in the context of a short course in advanced inventory control techniques. A class of eighteen interested and experienced people in the inventory control field met for two weeks during August 1968, for twenty class hours of instruction under the auspices of the Department of Computer and Information Science at the University of North Carolina in Chapel Hill. During the course the following topics, among others, were treated:

1. Exponential and weighted time series smoothing for demand forecasting.
2. Probability distributions of demand.
3. Evaluation of order-quantity computations.
4. Variable safety level concepts.
5. Performance budgeting.
6. Master record file design.

Near the end of the course, each participant was given, as examination problems, the statistics resulting from two different simulation runs. The participants were asked to rank the various management policies, month by month and to make other decisions regarding the policies, on the basis of the data presented. Each participant acting in the capacity of decision maker, or decider, using both printer output on one problem and graphic display presentation on the other, demonstrated his decision making ability.

The statistics from the simulation program were





displayed in a variety of formats on standard computer printer paper. The statistics included such data as percent availability of stock, number of purchase orders generated, lost sales and total dollar investment in inventory. The output was in the form of tabular listings and bar graphs. Also, for each problem, the IBM 2250 cathode ray tube display unit was programmed to show the same data as appear on the printer output, in fundamentally the same listing and graphical formats. Thus the experiment did not attempt to establish the relative effectiveness of listed versus plotted data. Instead it investigated the relative effectiveness of printed versus display presentation of such data. The display unit was under the control of the decider, who was able to specify which of the twenty-four months and the title of the displayed data which he wanted to see. In both display and printer output cases, he was expected to start at the first month and proceed through the months in order.

The course participants were divided into two test groups, each group consisting of nine individuals chosen at random. Tests were made in regard to:

1. Are decisions made faster? The elapsed clock time to decide on a ranking of policies which the decider would be willing to commit himself to for future action was tested.
2. Are decisions made earlier? The calendar month (out of the twenty-four months of simulation)



in which the decider feels he has enough data to commit himself to a ranking for future action was tested.

3. Are decisions made faster? The elapsed clock time to complete the remainder of the problem was tested. After committing himself to a ranking, the participant continued the problem through the twenty-fourth month.
4. Are decisions made more consistently? Individual correlation coefficients from the Spearman rank correlation test were computed.
5. To what degree do the members of each group agree among themselves in regard to the rankings? Kendall's W coefficient of concordance was computed for each cell of the Latin square.

In summary, the decider makes two basic kinds of decisions. He decides on a ranking at each month in turn, and he decides whether or not to indicate at this month that he feels confident enough of his ranking to predict that it will not significantly change over the remainder of the twenty-four months.

The resulting ordered lists of policy sets and the preparation times were analyzed in two basic ways. First, a 2 x 2 Latin square design with repeated measures was used in an analysis of variance. Rank order statistics were then used to test the consistency of each individual



decision maker and the concordance of a group of decision makers.

The statistical analysis of variance explored the differences between graphic and printed data presentation (treatment), between the first and second problems done by each participant (order), and isolated the differences between the performance of Group I and Group II.

The rank order of the policy sets at the decision month and at the end of the twenty-four month simulation was compared by the Spearman rank correlation coefficient test for each participant. These rank correlation coefficients were then arranged in the Latin square and an analysis of variance was performed. These tests determined whether treatment, order, or group composition caused differences in the consistency of a participant's final answer, after he had full data available, with that at the month of commitment to a policy ranking. A Kendall's coefficient of concordance was computed to measure how well the rankings produced by each test group on each problem at the decision month agreed among themselves.





## II. THE SIMULATION MODEL

This simulation program is a dynamic model of a general inventory control situation, and is intended to investigate inventory management policies in regard to forecasting, purchase quantity computation and service level over a range of items whose characteristics are defined by the user. The model was programmed in PL/I utilizing the IBM System 360 Model 75 at the Triangle Universities Computation Center, Research Triangle Park, North Carolina.

The model establishes identical initial conditions for each of the twelve policies and then generates a series of actions which represent daily business transactions under a particular policy. This series of daily actions is continued over a simulation cycle of two 360-day years for each policy in question. Reports are generated after each thirty-day month of simulation.

The Cartesian product of the sets of decision rules for service level (two rules), purchase quantity computation (three rules), and forecasting methods (two rules), gives twelve distinct cycles of 720 simulation days each.

The following are the formal properties of the simulation model. While the program is well documented





in the program listing, the tabulation below of the more important elements defines the model by a categorization appropriate to the simulation of economic systems.<sup>1</sup> We will define exogenous variables as those independent parameters which are treated by the system as inputs. They will, in general, be stochastic variables which are generated internally by the program's random variate generators. Status variables will describe the state of the system in relation to an appropriate time period, and are generally found to be in a feedback loop which interacts with the exogenous variables of a preceding time period. For instance, the on-hand inventory level, ONH, reflects a status at any one point in time, but is dependent on the values of several stochastic variables in preceding time periods. Endogenous variables will be the output result of interactions between exogenous and status variables. These are the numerical results which are found on the printed and cathode ray tube display output. Where status variables describe the system and interact with other variables, endogenous variables are collection points for output purposes only, and do not interact with other variables to affect system performance. Identities will be either definitions or tautologies which describe model components. For example, the status variable ONH

---

<sup>1</sup>Thomas H. Naylor, et al., Computer Simulation Techniques (New York: John Wiley & Sons, Inc., 1966), pp. 9-20, 159-173.



is defined by the identity  $ONH = ONH - TRANS + DUIN$ , where the identifier on the left at time  $t + 1$  is replaced by the expression on the right evaluated at time  $t$ . Parameters are those variables which define the inventory control policies and the distributions describing the environment under investigation, as distinguished from the exogenous variables that describe the particular experiences to which the policies are subjected. These the investigator changes in order to model his particular policies and simulated environments. In all cases the identifier mentioned is the one used in the PL/I program and is cross-listed in the Program Attribute Table.

#### EXOGENOUS VARIABLES

<u>Identifier</u>	<u>Description</u>
RAND	A uniformly distributed random variate in the range 0 to .99.
Y	A random variate from one of the ten random variate generators.
X	A number in the range one to ten which indicates which of the ten random variate generators (see page 20) is being called.

#### STATUS VARIABLES

<u>Identifier</u>	<u>Description</u>
ONH	On hand inventory; in units of issue per line item.
HQTY	Amount due in from purchase action initiated by a purchase quantity computation; in units of issue per line item.



#SOLD	Total units sold the previous year; in units of issue per line item.
RET	Total units returned by customers over the simulation cycle; in units of issue per line item.
FREQ	Frequency of satisfied demand this year; in number of demands satisfied per line item.
CRIT	Reorder point (variable safety level); in units of issue per line item.
TOTD	Total demand last year; in units of issue per line item.
CDF	Count-down variable which shows the number of days until a due-in order is expected to be received for this stock number; in days.
PRD	Forecast monthly demand; in units of issue per line item.
AUF	Annual usage (sales this year); in units of issue per line item.
AND	Total demand this year; in units of issue per line item.
NFREQ	Frequency of demands which could not be filled; in number of demands per line item.
MDEM	Demand this month; in units of issue per line item.
DEMD	The demand registered against a stock number on any particular transaction; in units of issue per line item.
TRANS	The number sold to a customer as the result of a demand; in units of issue per line item.
BNPRICE	The computed cost price of on-hand inventory after receipt of due-in merchandise; in dollars per unit of issue.





NEX	The mean of the number of units returned by a customer against a stock number; in units of issue per line item.
FCST1	Forecast demand at previous time period; in units of issue per line item.
TRND	Demand trend; in units of issue per line item.
ERROR	Error in forecast of demand during last time period; in units of issue per line item.
TREND	Smoothed trend forecast; in units of issue per line item.
EXPCDMD	Expected demand after trend correction to forecast demand; in units of issue per line item.
MAD	Mean absolute deviation of true demand from expected demand; in units of issue per line item.
MRD	Maximum reasonable demand from variable safety level computations; in units of issue per line item.

#### ENDOGENOUS VARIABLES

<u>Identifier</u>	<u>Description</u>
AVAIL	Availability, the ratio of units of satisfied demand to units of total demand; in percent per policy.
AV\$INV	Total inventory investment, the cost of the on-hand inventory plus due-in; in dollars per policy.
LOSS	Lost sales, the cost of not having stock which was demanded over the life of the simulation; in dollars per policy.
PURCH	Cumulative number of purchase actions generated; in number of purchase actions per policy.



1. The first part of the report deals with the general situation of the country and the progress of the work during the year.

2. The second part deals with the results of the work done during the year and the progress of the work during the year.

3. The third part deals with the results of the work done during the year and the progress of the work during the year.

4. The fourth part deals with the results of the work done during the year and the progress of the work during the year.

5. The fifth part deals with the results of the work done during the year and the progress of the work during the year.

6. The sixth part deals with the results of the work done during the year and the progress of the work during the year.

7. The seventh part deals with the results of the work done during the year and the progress of the work during the year.

8. The eighth part deals with the results of the work done during the year and the progress of the work during the year.

9. The ninth part deals with the results of the work done during the year and the progress of the work during the year.

10. The tenth part deals with the results of the work done during the year and the progress of the work during the year.

11. The eleventh part deals with the results of the work done during the year and the progress of the work during the year.

12.

13.

14.

15.

16.

17.

18.

19.

20.

21.

22.

23.

24.

S_CLITEM (1)	Total of items with unit price less than \$25.00 and annual demand less than \$100.00; in line items per policy.
S_CLITEM (2)	Total of items with unit price less than \$25.00 and annual demand $\geq$ \$100.00 and $\leq$ \$1,000; in line items per policy.
S_CLITEM (3)	Total of items with unit price $<$ \$25.00 and annual demand $>$ \$1,000; in line items per policy.
S_CLITEM (4)	Total of items with unit price over \$25.00; in line items per policy.
S_LNITEM (1)	Total of items with availability $<$ 70 percent; in line items per policy.
S_LNITEM (2)	Total of items with availability $\geq$ 70 percent to $<$ 75 percent; in line items per policy.
S_LNITEM (3)	Total of items with availability $\geq$ 75 percent to $<$ 80 percent; in line items per policy.
S_LNITEM (4)	Total of items with availability $\geq$ 80 percent to $<$ 85 percent; in line items per policy.
S_LNITEM (5)	Total of items with availability $\geq$ 85 percent to $<$ 90 percent; in line items per policy.
S_LNITEM (6)	Total of items with availability $\geq$ 90 percent to $<$ 95 percent; in line items per policy.
S_LNITEM (7)	Total of items with availability $\geq$ 95 percent; in line items per policy.

### IDENTITIES

The following identities are expressed in terms of a programming language assignment statement, where the identifier on the left at time  $t + 1$  is replaced by the



expression on the right, evaluated at time  $t$ .

TRANS = DMD if total demand satisfied, else  
TRANS = ONH; in units of issue per line  
item.

#SOLD =  $\sum_{i=1}^{BM} \text{TRANS}$ ; in units of issue per line  
item.

BNPRICE =  $\frac{\text{ONH} * \text{PRICE} + \text{BQ} * \text{DUIN}}{\text{ONH} + \text{DUIN}}$ ; in dollars per  
line item.

PRICE = BNPRICE; in dollars per line item.

ONH = ONH - TRANS + DUIN; in units of issue  
per line item.

DUIN = HQTY + (TIME/30) \* PRD; in units of  
issue per line item.

PRD = PRD \* (1 - BCONSTANT) + MDEM \*  
BCONSTANT; in units of issue per line  
item.

HQTY =  $\sqrt{\frac{2 * \text{HORDERCOST} * \text{AUF}}{\text{HHOLD COST}}}$  for EOQ; in units  
of issue per line item. (Other  
decision rules for HQTY are outlined  
in text.)

EXPEC = AND/(FREQ + NFREQ); in units of issue  
per line item.

STDV = EXPEC/2; in units of issue per line  
item.

NEX = AND/FREQ; in units of issue per line  
item.

MDEM =  $\sum_{BD=1}^{30} \text{DEMD}$ ; in units of issue per line  
item.

STDV = NEX/10; in units of issue per line  
item.

TRND = PRD - FCST1; in units of issue per line  
item.

TREND = (BCONSTANT \* TRND) + (1 - BCONSTANT) \*  
TREND; in units of issue per line  
item.



ERROR = MDEM - EXPCDMD; in units of issue per line item.  
 EXPCDMD =  $\text{PRD} + \text{TREND} * (1 - \text{BCONSTANT}) / \text{BCONSTANT}$ ; in units of issue per line item.  
 MAD =  $\text{ABS}(\text{BCONSTANT} * \text{ERROR}) + (1 - \text{BCONSTANT}) * \text{MAD}$ ; in units of issue per line item.  
 MRD =  $( (\text{TIME}/30) * \text{EXPCDMD} ) + \text{KFACT} * \text{MAD} * \text{SQRT}(\text{TIME}/30)$ ; in units of issue per line item.

### PARAMETERS

<u>Identifier</u>	<u>Description</u>
PRICE	Cost price of inventory item; in dollars per unit of issue.
TIME	Purchase lead time; in days per purchase order per line item.
BCONSTANT	Exponential smoothing constant; a decimal fraction.
BSAFE	Fraction of demand unsatisfied, for use in variable safety level computation; a decimal fraction.
BYRC	Number of years in total time period of simulation; in years per simulation cycle.
BD	Number of days per month; a numeric constant in the range one to thirty.
HORDERCOST	Order cost; in dollars per purchase action.
HHOLDCOST	Holding cost; in dollars per year.
HMAX	Maximum purchase quantity; in units of issue per transaction.
HMIN	Minimum purchase quantity; in units of issue per transaction.



ANNUAL\$DMD

Annual dollar demand comparand  
for purchase quantity compu-  
tation; in dollars per year.

In order to provide for future demands against inventory, it is usual to use past demand history as a basis for forecasting this future demand. In this model, exponential smoothing<sup>2</sup> was used to operate on past and present demand to produce a forecast demand for use in purchase quantity and safety level computation. Single exponential smoothing is defined to be:

$$Y_t = \alpha X_t + (1 - \alpha) Y_{t-1}$$

where  $Y_t$  is the average computed for the end of time period  $t$  and  $X_t$  is the demand registered in time period  $t$ . The smoothing process can be recursively applied, yielding higher degrees of smoothing expressed by:

$$Y_t^k = \alpha Y_t^{k-1} + (1 - \alpha) Y_{t-1}^k$$

where  $k$  is a superscript which defines the degree of smoothing, and  $t$  indicates the time period in question. Double smoothing, for example, requires the following calculations:

$$Y_t^1 = \alpha X_t + (1 - \alpha) Y_{t-1}^1$$

$$Y_t^2 = \alpha Y_t^1 + (1 - \alpha) Y_{t-1}^2$$

While of interest for theoretical purposes, high orders of smoothing in general provide little practical

---

<sup>2</sup>Robert G. Brown, Statistical Forecasting for Inventory Control (New York: McGraw-Hill Book Company, 1959), pp. 45-51.





benefit in terms of forecast accuracy, and obviously increase file storage space requirements and computation time. For these reasons, only single and double smoothing have been used in the model.

From the definition, however, it can be seen to be a trivial matter in PL/I to program a recursive function call for any desired degree, assuming the availability of  $k$  storage locations with each master file stock item for  $k^{\text{th}}$  degree exponential smoothing to hold the  $k$  values of  $y_k$ .

The specification of two alternative values for the  $\alpha$ -constant in the forecasting routine is at the option of the user. For preliminary investigation, values of 0.2 and 0.4 are customary, but, of course, any values may be specified. The specification of the  $\alpha$ -constants completes the first of the three sets of parameters needed to identify the twelve policies to be simulated.

The second set of parameters controls the purchase quantity computation. Here three choices are available for composing the twelve-policy set and the user may either use the three decision rules embedded in the program, specifying their parameters as desired, or entirely new rules may be specified in any combination with little additional effort.

When the stock on-hand gets to a certain level (CRIT) a computation is made to determine how much to purchase to replenish this stock. This aperiodic action



was taken in lieu of a periodic reorder cycle on the assumption that we are dealing with high demand items (as opposed to insurance items) and that if there is new information, use should be made of it. This new information is in the form of the variable safety level (CRIT) and the total demand forecast. This philosophy could have been carried one step further and a new exponentially smoothed demand produced at this point, but with monthly forecasting in use this was deemed unnecessary. The purchase quantity computation methods were chosen in order to provide three quite distinguishable computations so as to give clear alternatives in the evaluation of the results. The purchase quantity computation methods chosen to model were:

1. Standard EOQ computation<sup>3</sup> with:

HORDERCOST = Order cost = \$70.

HHOLD COST = Holding cost = 15 percent of price.

AUF = Annual demand from master file.

HMAX = Maximum buy = twelve months of forecast demand.

HMIN = Minimum buy = one month of forecast demand.

2. Modified EOQ computation with different parameters in order to accentuate any distinction with the standard EOQ computation:

---

<sup>3</sup>G. W. Plossl and O. W. Wight, Production and Inventory Control (Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1967), pp. 388-390.



HORDERCOST = Order cost = \$100.

HMAX = Maximum buy = twenty-four months forecast demand.

HOLDCOST = Holding cost = 15 percent of price.

HMIN = Minimum buy = two months forecast demand.

The decision rule for the modified EOQ computation was to use the standard EOQ formula with the above parameters if the forecast annual demand in dollars was less than \$1,000. If the annual demand forecast was equal to or greater than \$1,000, but less than \$10,000, the purchase was to be simply twelve months of predicted demand. Otherwise the standard EOQ formula was used with the above parameters except that HORDERCOST = \$500.

3. Three way function of annual dollar demand forecast:

1) If forecast dollar demand is less than \$1,000, buy twenty-four months of forecast demand.

2) If forecast dollar demand is equal to or greater than \$10,000, buy two months of forecast demand.

3) Otherwise buy six months of forecast demand.

These three methods represent enough of a variation so that the results show significant differences from which to choose. Of course any other set of purchase quantity computation methods could have been chosen. Since the number of simulation cycle runs is dependent on the





Cartesian product of the overall number of rules in each set under consideration, the addition of more rules must be considered with total run time in mind.

The third and last set of rules involves a choice of two performance levels or service levels.<sup>4</sup> The service level is in essence the desired availability of stock in the system. For instance, a service level of 90 percent indicates that the system economics are expected to be able to satisfy 90 percent of the demand against system items. The parameter BSAFE may be set to reflect two different service levels which are to be investigated. If service level is defined as fraction of demand to be satisfied from stock, BSAFE will be defined as fraction of demand unsatisfied, or fraction of demand back ordered. If we are investigating the feasibility of system support for service levels of 95 percent and 98 percent, the values of the parameter BSAFE would be specified as 0.05 and 0.02. The principal use of this parameter is in the computation of the variable safety level,<sup>5</sup> CRIT. This safety level, or reorder point, varies with each item as the item's demand pattern changes. The variable safety level computation takes into account past forecast errors and produces a small safety level of stock for a well behaved demand pattern and a large safety level for a demand pattern

---

<sup>4</sup>Brown, Statistical Forecasting, pp. 105-116.

<sup>5</sup>Ibid., pp. 136-159.





which produces large forecast errors. The variable safety level computation is accomplished monthly, at forecast time, for each stock item in the file.

For the purposes of this model, an inventory control policy was defined to be one of the twelve possible combinations made by selecting one element from each of the above three sets of parameters. For example, one policy might be  $\alpha = .15$ , standard EOQ with parameters as given in the model, BSAFE = .02.

The user has the ability to specify initial stock conditions and also the probability distributions for the following variables:

1. Return of material to stock.
2. Frequency of demand.
3. Price change on receipt of material.
4. Quantity demanded against individual items.

In addition, the user may wish to simulate other types of transactions than are provided for in the model. The program modification in this case would be left to the user. To allow flexibility in introducing new subroutines to extend or change the model, the following pseudo-random variate generators are available as simple PL/I function calls:

1. Uniform.
2. Normal.
3. Poisson.
4. Geometric.



5. Gamma.
6. Binomial.
7. Hypergeometric.
8. Log-Normal.
9. Negative Binomial.
10. Exponential.

The index of the random variate generator in the above list corresponds to the value of the exogenous variable X (see page 9 ).



### III. THE PRESENTATION TO THE DECISION MAKER

The simulation statistics that are available on a monthly basis are shown in Figures 2, 3, and 4. Figure 1 is a three-dimensional representation of how data are collapsed into the matrix in Figure 2. The matrix entries correspond to the number of line items which fell into that matrix element as of that point in time. The vertical columns, or categories, are totaled to show the number of line items in each price-demand category. The totals at the bottom of the figure are sub-tabulated by the price-demand categories. Only the total investment figure is scaled. A scale of 100,000 was applied to this due to restrictions on the allowable number of characters per line on the display surface of the cathode ray tube. A total investment of \$423,251.81 would, therefore, be reflected as \$4.23.

Figure 3 is a tabular monthly recapitulation which is ordered by several of the statistics of interest. This is the counterpart of Figure 4 which shows one statistic across all policies as a bar graph. The printed bar graphs are produced monthly for lost sales, purchase order generation, total investment and stock availability.

Once the printer output was specified, it was natural to attempt to reproduce this presentation on



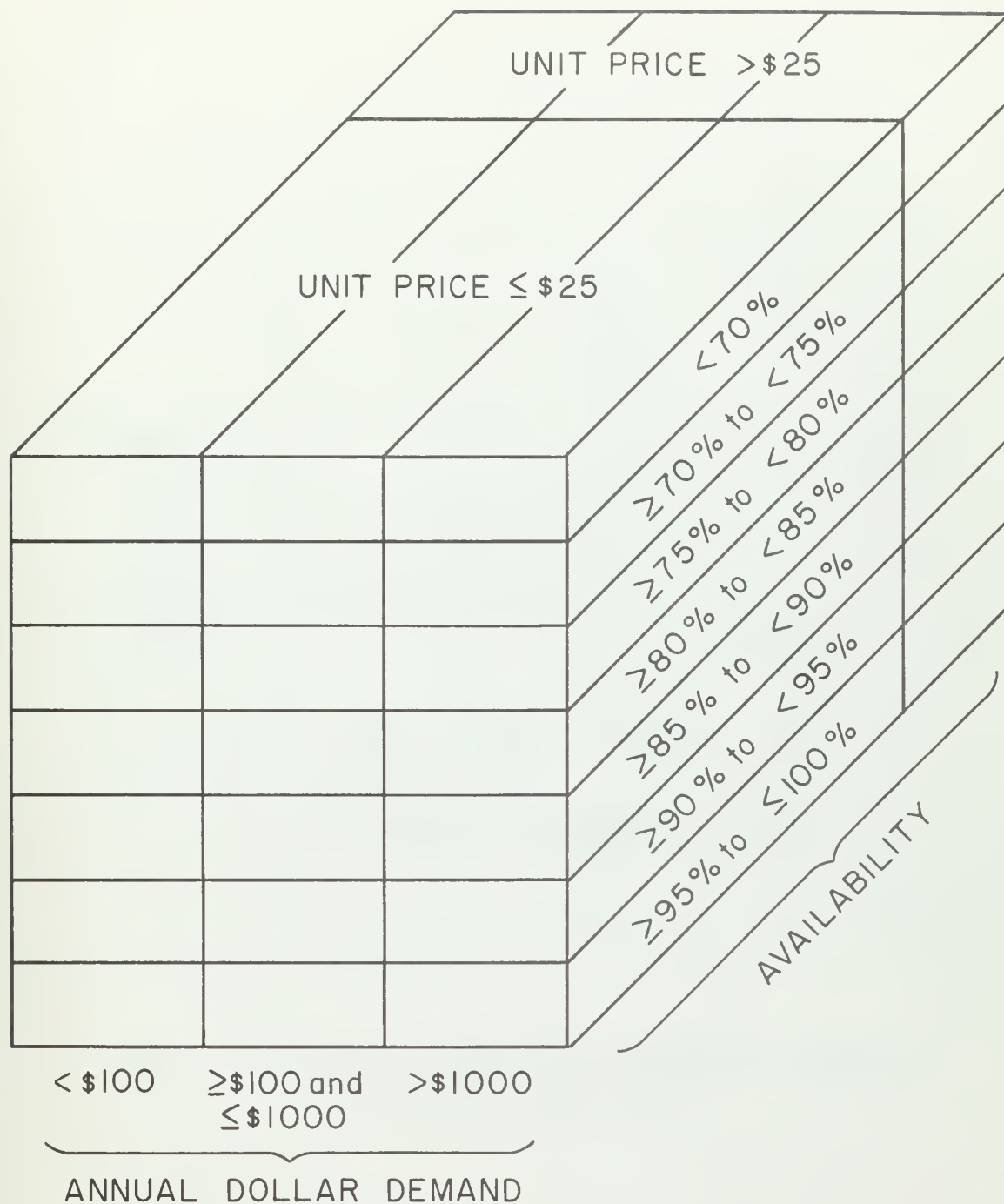


Fig. 1.--Three-dimensional representation of Figure 2





MONTH 16

POLICY 2

PROBLEM # 13

ANNUAL DOLLAR DEMAND AVAILABILITY	UNIT PRICE < \$25			UNIT PRICE > \$25		LINE TOTALS
	< \$100	≥ \$100 & ≤ \$1,000	> \$1,000			
< 70%	0	0	0	0		0
≥ 70% to < 75%	0	0	0	0		0
≥ 75% to < 80%	0	0	0	1		1
≥ 80% to < 85%	0	0	2	0		2
≥ 85% to < 90%	0	0	0	3		3
≥ 90% to < 95%	0	0	1	1		2
≥ 95% to ≤ 100%	5	4	13	7		29
CATEGORY TOTALS:	5	4	16	12	TOTAL NUMBER OF ITEMS:	37

CATEGORY	LOST SALES	% AVAILABILITY	TOTAL INVESTMENT
1	\$0.00	99.99	\$0.01
2	\$6.78	99.23	\$0.02
3	\$16,583.44	99.97	\$1.07
4	\$125,480.56	99.69	\$3.46
OVERALL:	\$142,070.75	92.90	\$4.57

CUMULATIVE NUMBER OF PURCHASE ACTIONS FOR THIS POLICY = 57  
 CUMULATIVE NUMBER OF TRANSACTIONS GENERATED UNDER THIS POLICY = 1931

Fig. 2.--Detailed listing of simulation results  
by policy within month



RESULTS OF SIMULATION BY : POLICY NUMBER					MONTHS 16		PROBLEM 6.13	
POLICY NUMBER	AVAILABILITY	TOTAL INVESTMENT	LOST SALES	PURCHASE ACTIONS	TRANSACTIONS			
1	89.37	\$3.40	\$222,191.56	108	1937			
2	92.90	\$4.57	\$142,070.81	57	1931			
3	89.73	\$2.73	\$50,025.95	87	1944			
4	89.32	\$2.78	\$153,479.62	73	1896			
5	92.98	\$4.79	\$232,559.37	65	1956			
6	91.01	\$3.40	\$224,220.50	89	1909			
7	95.15	\$2.77	\$23,692.03	73	1933			
8	96.31	\$4.27	\$42,522.65	56	1926			
9	94.60	\$4.16	\$145,429.12	98	1881			
10	86.68	\$2.60	\$144,235.37	71	1889			
11	93.78	\$5.37	\$152,888.06	53	2018			
12	93.46	\$2.86	\$85,199.18	61	1922			
RESULTS OF SIMULATION BY : TOTAL INVESTMENT								
POLICY NUMBER	AVAILABILITY	TOTAL INVESTMENT	LOST SALES	PURCHASE ACTIONS	TRANSACTIONS			
10	86.68	\$2.60	\$144,235.37	71	1889			
3	89.73	\$2.73	\$50,025.95	87	1944			
7	95.15	\$2.77	\$23,692.03	73	1933			
4	89.32	\$2.78	\$153,479.62	73	1896			
12	93.46	\$2.86	\$85,199.18	61	1922			
1	89.37	\$3.40	\$222,191.56	108	1937			
6	91.01	\$3.40	\$224,220.50	89	1909			
9	94.60	\$4.16	\$145,429.12	98	1881			
8	96.31	\$4.27	\$42,522.65	56	1926			
2	92.90	\$4.57	\$142,070.81	57	1931			
5	92.98	\$4.79	\$232,559.37	65	1956			
11	93.78	\$5.37	\$152,888.06	53	2018			
RESULTS OF SIMULATION BY : AVAILABILITY								
POLICY NUMBER	AVAILABILITY	TOTAL INVESTMENT	LOST SALES	PURCHASE ACTIONS	TRANSACTIONS			
8	96.31	\$4.27	\$42,522.65	56	1926			
7	95.15	\$2.77	\$23,692.03	73	1933			
9	94.60	\$4.16	\$145,429.12	98	1881			
11	93.78	\$5.37	\$152,888.06	53	2018			
12	93.46	\$2.86	\$85,199.18	61	1922			
5	92.98	\$4.79	\$232,559.37	65	1956			
2	92.90	\$4.57	\$142,070.81	57	1931			
6	91.01	\$3.40	\$224,220.50	89	1909			
3	89.73	\$2.73	\$50,025.95	87	1944			
1	89.37	\$3.40	\$222,191.56	108	1937			
4	89.32	\$2.78	\$153,479.62	73	1896			
10	86.68	\$2.60	\$144,235.37	71	1889			
RESULTS OF SIMULATION BY : LOST SALES COSTS								
POLICY NUMBER	AVAILABILITY	TOTAL INVESTMENT	LOST SALES	PURCHASE ACTIONS	TRANSACTIONS			
7	95.15	\$2.77	\$23,692.03	73	1933			
8	96.31	\$4.27	\$42,522.65	56	1926			
3	89.73	\$2.73	\$50,025.95	87	1944			
12	93.46	\$2.86	\$85,199.18	61	1922			
2	92.90	\$4.57	\$142,070.81	57	1931			
10	86.68	\$2.60	\$144,235.37	71	1889			
9	94.60	\$4.16	\$145,429.12	98	1881			
11	93.78	\$5.37	\$152,888.06	53	2018			
4	89.32	\$2.78	\$153,479.62	73	1896			
1	89.37	\$3.40	\$222,191.56	108	1937			
6	91.01	\$3.40	\$224,220.50	89	1909			
5	92.98	\$4.79	\$232,559.37	65	1956			

Fig. 3.--Sorted listing of simulation results  
by policies within month



[illegible]

Fig. 4.--Bar graph of simulation results by month across all policies



another device for the experiment. The experimental machine was chosen to be the IBM 2250 Model 1 cathode ray tube display device.<sup>6</sup> See Figure 5. This particular device is program controlled from the multiprogramming environment of the IBM System 360 Model 40 computer at the University of North Carolina Computation Center. A program was written to take the data output of the simulation program and mimic the printer output formats on the display device. Some exploration of the power and utility of such a display device was made within the confines of the desirability of maintaining similar presentation format for experimental design purposes. Figures 6 and 7 are the display device representations of Figures 2 and 3 and show some rearrangement of material due to display character and line number limitations.<sup>7</sup> Figure 8 shows the first marked departure from the printer output of Figure 4. The layout is essentially the same, but the presentation is dynamic. A programmed timer was specified in the display program to advance all lines of the bar graph simultaneously at one second intervals. If the user specifies a month at which he wishes to examine the data, and then specifies the bar

---

<sup>6</sup>International Business Machines Corporation, IBM System/360 Component Description, IBM 2250 Display Unit Model 1, Form A27-2701-1 (January 27, 1967), pp. 5-37.

<sup>7</sup>International Business Machines Corporation, IBM System/360 Operating System, Graphic Programming Services for IBM 2250 Display Unit, Form C27-6909-4 (December 1967), p. 14.







Fig. 5.--IBM 2250 cathode ray tube display device



MONTH 16

POLICY 02

PROBLEM # 13

	UNIT PRICE < \$ 25-----> -----AND ANNUAL DOLLAR DEMAND----->			UNIT PRICE > \$ 25	ROW TOTALS
	< \$ 100	>=100 TO 1000	>\$1000		
<70%	0	0	0	0	0
>=70 TO <75%	0	0	0	0	0
>=75 TO <80%	0	0	0	1	1
>=80 TO <85%	0	0	2	0	2
>=85 TO <90%	0	0	0	3	3
>=90 TO <95%	0	0	1	1	2
>=95 TO 100%	5	4	13	7	29
CATEGORY TOTALS:	5	4	16	12	
TOTAL NUMBER OF ITEMS:					37

CATEGORY	LOST SALES	% AVAILABLE	TOTAL INVESTMENT
1	0	99	\$ .01
2	6	99	\$ .02
3	16,583	93	\$ 1.07
4	125,480	88	\$ 3.46
OVERALL	142,070	92	4.57

CUMULATIVE NUMBER OF PURCHASE ACTIONS FOR THIS POLICY = 57

Fig. 6.--Display of Figure 2



MONTH 16 POLICY	% AVAILABLE	PROBLEM 13 LOST SALES	TOT INVSTMT	PURCHASES
1	89	\$ 222,191	\$ 3.40	108
2	92	\$ 142,070	\$ 4.57	57
3	89	\$ 50,025	\$ 2.73	87
4	89	\$ 153,479	\$ 2.78	73
5	92	\$ 232,559	\$ 4.79	65
6	91	\$ 224,220	\$ 3.48	89
7	95	\$ 23,691	\$ 2.77	73
8	96	\$ 42,522	\$ 4.27	56
9	94	\$ 145,429	\$ 4.16	98
10	86	\$ 144,235	\$ 2.60	71
11	93	\$ 152,887	\$ 5.37	53
12	93	\$ 85,199	\$ 2.86	61

-----RESULTS SORTED BY MAJOR CATEGORY :-----

POLICY %AVAILABLE	POLICY	LOST SALES	POL. INVST.	POL. PURCH.
8 96	7	23,691	10 \$ 2.60	11 53
7 95	8	42,522	3 \$ 2.73	8 56
9 94	3	50,025	7 \$ 2.77	2 57
11 93	12	85,199	4 \$ 2.78	12 61
12 93	2	142,070	12 \$ 2.86	5 65
5 92	10	144,235	1 \$ 3.40	10 71
2 92	9	145,429	6 \$ 3.48	4 73
6 91	11	152,887	9 \$ 4.16	7 73
3 89	4	153,479	8 \$ 4.27	3 87
1 89	1	222,191	2 \$ 4.57	6 89
4 89	6	224,220	5 \$ 4.79	9 98
10 86	5	232,559	11 \$ 5.37	1 108

Fig. 7.--Display of Figure 3









graph he wishes to see, the graph begins at month 1 and then advances at one second intervals to the month requested. The user may also specify half-second intervals by a light pen detect. It is intended that this movement of the bar graph give a feeling for the history and current derivative of the statistic under scrutiny as well as the static numeric results of the simulation. In addition to the movement of the bar graph, the symbols  $>$ ,  $<$  and  $=$  are used on the display unit to indicate whether the value of a particular bar has increased, decreased or remained equal in comparison with the previous month.

Figure 9 has no printed correspondent in this experiment. While the data for cost of lost sales, availability, total dollar investment and number of purchase actions generated are available as individual bar graphs as represented by Figure 8, all four of these are available at the same time in the Quadrant Graph. All move simultaneously under the same program timer control as the individual graphs. Each graph is scaled uniquely and, of course, represents the same data as are available on the individual graphs, but with the obvious reduction in resolution since each line of a graph has only approximately half the width of the display device screen to represent the data as does the individual graph.

Figure 10 depicts the Program Function Keyboard (PFK) which is the means by which the user indicates to the computer what he wishes to see on the display device.





Fig. 9.--Quadrant graph of simulation results  
by month across all policies



Overlay Code

0 1 2 3 4 5 6 7

<u>INVENTORY</u> <u>CONTROL PROGRAM</u>		○		<u>J. S. PROKOP</u>	
<u>Instructions/</u> <u>Termination</u>		<u>Month/</u> <u>Policy</u>	<u>Month/</u> <u>Policy</u>	<u>Month/</u> <u>Policy</u>	
0	○	1	○	2	○
<u>Month/</u> <u>Policy</u>	<u>Month/</u> <u>Policy</u>	<u>Month/</u> <u>Policy</u>	<u>Month/</u> <u>Policy</u>	<u>Month/</u> <u>Policy</u>	<u>Month/</u> <u>Policy</u>
4	○	5	○	6	○
<u>Month/</u> <u>Policy</u>	<u>Month/</u> <u>Policy</u>	<u>Month/</u> <u>Policy</u>	<u>Month</u>	<u>Month</u>	<u>Month</u>
10	○	11	○	12	○
<u>Month</u>	<u>Month</u>	<u>Month</u>	<u>Month</u>	<u>Month</u>	<u>Month</u>
16	○	17	○	18	○
<u>Month</u>	<u>Month</u>	<u>Month</u>	<u>Sorted</u> <u>Listing</u>	<u>Availab.</u> <u>Graph</u>	<u>Logst Sales</u> <u>Graph</u>
22	○	23	○	24	○
	<u>Purchase</u> <u>Graph</u>	<u>Investmt</u> <u>Graph</u>	<u>Advance</u>	<u>Quadrant</u> <u>Graph</u>	
28	○	29	○	30	○
				31	○

Fig. 10.---Programmed function keyboard



Key 0 will always display the user instructions if the plastic overlay, with no sense switch cut-outs,<sup>8</sup> is in place. Otherwise, key 0 will return control to the graphics monitor.

There is a two-step procedure for obtaining a display:

1. Indicate which month is to be seen.
2. Indicate the display data desired.

In order to prevent inadvertent display of data beyond the month desired, month change was set up as a two-step action. First, key 30 must be depressed (at which time a tone signal is heard), then the key corresponding to the month is depressed. This key is lighted by program control as a reminder to the user of his simulated month of furthest advance. At this time, no change is noted on the display device screen. When a key corresponding to a display is depressed, the display then changes. Keys 1-12 will display Figure 6 for policies 1-12 for the month indicated. Keys 13-24 will be ignored by the program at this point. Note that keys 1-24 signify different events depending on whether they are immediately preceded by a key 30 depression or not. Key 25 will produce Figure 7, keys 26-29 will produce Figure 8 for Availability, Cost of Lost Sales, Purchase Actions Generated, and Total Dollar Investment respectively. Key 31 will produce Figure 9.

---

<sup>8</sup>IBM, Component Description, pp. 30-32.





The user is allowed to refer to any past month's data. He may not, however, move ahead in simulated time by more than one month at a time.



#### IV. THE SUBJECTS OF THE EXPERIMENT

The experiment was designed to examine the decision-making processes of experienced, practical and interested professional administrators of inventory control, to find out if graphical display devices do, in fact, assist the decision-making process. The subjects were solicited in such a way as to insure a professional audience. The short course was designed as a graduate-level course which would attract only those who were interested in the subject matter and who were prepared to understand the material. The simulation model was the experimental vehicle and was used in examples and problems which were integrated into the course. The participants came to the course in statistical inventory control for their own professional advancement, under the auspices of the firms for which they worked. The participants were not informed of the experiment which was being conducted, and they received the full measure of instruction for which they came. The control of the human factors in the experiment proceeded quite smoothly with nothing to indicate bias or data contaminated by the fact that an experiment was in progress.

The individuals who comprised the experimental group were solicited from nearby major manufacturing concerns specifically for the short course in statistical



inventory control. It was requested that candidates for the course have experience in inventory control and be in a decision-making position in the company. No more than two candidates were to be sponsored by any one company.

The following firms sent people to the course:

Burlington Hosiery, Burlington

Burlington House Division, Burlington

Cardinal Products, Durham

Coca Cola, Durham

Golden Belt Manufacturing Company, Durham

International Business Machines, Raleigh

Kayser-Roth Hosiery, Burlington

Liggett and Meyers, Durham

Sears Roebuck, Greensboro

Sperry Rand, Durham

Virginia Mills, Swepsonville

Westinghouse Electric, Raleigh

The participants attended the course between the hours of 9:00 and 11:00 A.M. weekdays for a two-week period. The morning hours were chosen specifically to insure that the attendees would have the active concurrence and support of their sponsoring firms and would be scheduled for this absence from their regular duties.

After a random division into two groups of nine individuals each, the following statistics applied to the individuals:



Individual	Years Experience in Inventory Control	Automated Inven- tory Control Experience	Presently in Decision- Making Capacity
<u>Group I</u>			
A	20	No	Yes
B	15	No	Yes
C	20	No	Yes
D	12	Yes	Yes
E	1	No	No
F	5	Yes	Yes
G	10	Yes	Yes
H	1	Yes	Yes
I	2	No	Yes
<u>Group II</u>			
J	5	Yes	Yes
K	1	No	No
L	5	Yes	Yes
M	0	No	No
N	2	No	Yes
O	20	Yes	Yes
P	6	Yes	Yes
Q	2	No	Yes
R	2	No	Yes

After a substantial part of the course material had been given, the participants were given either a problem



TABLE I			
Year	Population	Area	Population Density
1950	1,000,000	100,000	10
1960	1,500,000	150,000	10
1970	2,000,000	200,000	10
1980	2,500,000	250,000	10
1990	3,000,000	300,000	10
2000	3,500,000	350,000	10
2010	4,000,000	400,000	10
2020	4,500,000	450,000	10
2030	5,000,000	500,000	10
2040	5,500,000	550,000	10
2050	6,000,000	600,000	10
2060	6,500,000	650,000	10
2070	7,000,000	700,000	10
2080	7,500,000	750,000	10
2090	8,000,000	800,000	10
2100	8,500,000	850,000	10

TABLE II			
Year	Population	Area	Population Density
1950	1,000,000	100,000	10
1960	1,500,000	150,000	10
1970	2,000,000	200,000	10
1980	2,500,000	250,000	10
1990	3,000,000	300,000	10
2000	3,500,000	350,000	10
2010	4,000,000	400,000	10
2020	4,500,000	450,000	10
2030	5,000,000	500,000	10
2040	5,500,000	550,000	10
2050	6,000,000	600,000	10
2060	6,500,000	650,000	10
2070	7,000,000	700,000	10
2080	7,500,000	750,000	10
2090	8,000,000	800,000	10
2100	8,500,000	850,000	10

NOTE: The data in this table are based on the 1950-2010 period. The population density is calculated as the population divided by the area. The area is calculated as the population divided by the population density.

book with twenty-four months of simulation output (Figures 2, 3, and 4) or assigned to work a problem on the display unit at individual laboratory sessions. The second problem was assigned later in a similar manner except the groups now used the presentation method that they had not used on the first problem. It was explained that congestion on the display unit precluded everyone working both problems in this way. There was no attempt made to identify individuals by group membership or emphasize this distinction.

The participants were not graded on these problems and were encouraged to work them as part of the education process of the course to investigate some of the fundamental properties of the inventory control policies. Although all of the inventory control policies simulated for the two problems were discussed in class, the participants were not given the correspondence to identify specific policies by number on the problem to avoid obvious bias. The participants, then, were asked to make their judgment based solely on the evidence without subjecting them to the prejudice of prejudgment. The two problems were of equivalent difficulty, the only distinction being the order in which they were presented.

An operational environment was explained for the problem. Higher level management had presented the participant with the output data for the set of twelve inventory control policies, and had requested a recommended ranking so that an implementation of policies could be decided



upon. The participant was encouraged by management to present his recommended ranking as soon as possible, but cautioned that implementation of the recommended policies could have serious repercussions in the firm, so his best professional judgment was required. Figure 11 was the work sheet to be used. Participants were told to ignore the first two items in the top line as they had no relevance in this problem (Group #, USING:). Participants were reminded to keep accurate figures on elapsed time spent studying each month's data, and were also reminded to mark the decision month at which they would have presented that ranking as a recommendation to higher level management, prior to completing the rest of the form.

In this experiment there were two possible approaches to the problem of evaluating the quality of the decisions made. The first and most commonly used approach is to establish and furnish the subjects with the criteria of excellence which is to be used in evaluating the decisions. For example, in many cases the subjects are first given instruction, then tested to ascertain how well they perform as measured against the pre-established criteria.

This method of measurement may be dependable when the material is objective and the criteria are easily established. In inventory control (and many other problems) the weightings to be used in reaching a decision are highly individual. For example, one firm may emphasize









high availability, whereas another may give heavy weight to low investment. When the experimental subjects are experienced in the field and have developed their own criteria for decisions, the method of measuring against instructor-set criteria is unsuitable. One cannot know the extent to which the subjects followed the instructor's criteria and the extent to which they followed their own.

A distinguishing feature of this experiment is that it took an alternative approach to the problem of evaluating the quality of decisions. The subjects were well acclimated to the decision process by experience, and were familiar with and interested in the substance of the course. Therefore, they were given no weightings, no criteria of excellence by the instructor. Instead they were explicitly told to apply their own several diverse sets of criteria. The quality of their decisions was then measured by the consistency of each subject's own ranking at the month of commitment with the later ranking when full information was available.

Although this second approach diminishes the meaningfulness of comparisons which show how well the subjects agree with each other, such measures were taken as a matter of interest. The use of self-consistency rather than artificial criteria does, however, improve the credibility of the other statistics. It also substantially improves the generalizability of the results, for it shows the effect on decision makers when deciding by their own





standards, rather than by those dictated by a simplified theoretical model. In substance, we have separated the decision process from the classroom environment and inserted it into an operational environment.



## V. THE DESIGN OF THE EXPERIMENT

Two main analyses were performed on the data: an analysis of variance of the Latin square design, and a computation of rank order statistics. We will first turn our attention to the Latin square design.

The concept of a generalized Latin square design is traditionally introduced by explaining a checkerboard pattern of equal-area plots of land used in agricultural experiments. If we are investigating the effect of soil treatment on the yield of a crop, the plots are chosen at random in such a way that every row and every column is a complete replication, which is to say that each treatment will occur once and only once in each row and each column on the checkerboard. In an experiment, this design tends to avoid any systematic variations in effects, such as local soil variations, which are not due to the treatment.

In the Latin square design of this experiment, it should be noted that instead of experimenting on different subjects in each cell of the square, the same experiment group was involved in both cells of a row. In effect, each group acted as its own control group. Experiments in which the same subjects are used under all  $q$  treatments require  $q$  observations on each subject, and are called experiments with repeated measures. In this experiment



$q = 2$  since we are dealing with two treatments and each subject is observed twice. This variation on the standard Latin square is described in detail on the following pages.

The formal design proceeds from a specification of the model to an analysis of variance and hence of the significance of observed differences. In the analysis of variance, the sum of squares (SS) is computed for rows, columns and treatments. From this, a variance estimate, or mean square (MS), is calculated for rows, columns and treatments. This is accomplished by dividing each sum of squares by the degrees of freedom (df) corresponding to that sum of squares. From here an F test<sup>9</sup> is used to find the significance level of the variation under investigation. We will be most concerned with the significance of any variance due to treatment.

	Order 1	Order 2
Group I (9 Individuals)	Display	Printer
Group II (9 individuals)	Printer	Display

Fig. 12.--Latin square design

The 2 x 2 Latin square with repeated measures<sup>10</sup>

---

<sup>9</sup>William L. Hays, Statistics for Psychologists (New York: Holt, Rinehart and Winston, 1963), pp. 368-369.

<sup>10</sup>B. J. Winer, Statistical Principles in Experimental Design (New York: McGraw-Hill Book Company, 1962), pp. 538-543.



of Figure 12 was used to analyze the data for:

1. Simulated calendar time to reach a decision (decision month),
2. Actual elapsed time to complete the rankings through the decision month,
3. Actual elapsed time to complete the problem after the decision month,
4. Correlation coefficients for the consistency of the committed decision with the final ranking,

in order to produce an analysis of variance of these statistics.

The particular design was chosen in part to isolate that variation due to natural differences between the experimental groups. As can be seen below, the separation of the source of variation due to group differences allows a better measure of that variation due to the form of data presentation, which is the statistic of real interest. An appropriate model for this 2 x 2 Latin square design is:<sup>11</sup>

$$E(X_{ijklm}) = \mu + \sigma_k + \pi_{m(k)} + \alpha_i + \beta_j \quad [1]$$

In [1],  $\mu$  is the grand mean,  $\sigma_k$  represents the effects associated with groups,  $\pi_{m(k)}$  the effects associated with subjects within groups,  $\alpha_i$  the effects due to the order in which each subject did the two experiments and  $\beta_j$  the effects due to the form of data presentation. The analysis and expected values of the mean

---

<sup>11</sup>Ibid., p. 539.





squares are shown in Figure 13, for the 2 x 2 case of the  $p \times p$  Latin square with  $n$  observations in each cell. Here  $A$  and  $B$  represent the factors associated with the  $\alpha$  and  $\beta$  effects.

Schematically the experiment may be represented by:

	$a_1$	$a_2$
$c_1$	$b_1$	$b_2$
$c_2$	$b_2$	$b_1$

Variation Source	df	$p=2, n=9$	E (MS)
<u>Between Subjects</u>	<u><math>np-1</math></u>	<u><math>= 17</math></u>	
Groups	$p-1$	$= 1$	$\sigma_{\epsilon}^2 + p\sigma_{\pi}^2 + np\sigma_{\delta}^2$
Subjects within Groups	$p(n-1)$	$= 16$	$\sigma_{\epsilon}^2 + p\sigma_{\pi}^2$
<u>Within Subjects</u>	<u><math>np(p-1)</math></u>	<u><math>= 18</math></u>	
A (Order)	$p-1$	$= 1$	$\sigma_{\epsilon}^2 + np\sigma_{\alpha}^2$
B (Treatment)	$p-1$	$= 1$	$\sigma_{\epsilon}^2 + np\sigma_{\beta}^2$
Error (Within)	$p(n-1)(p-1)$	$= 16$	$\sigma_{\epsilon}^2$

Fig. 13.--Analysis of the 2 x 2 Latin square design

The computational procedure for the analysis of variance involves a straightforward partition of the total sum of squares.<sup>12</sup> For this particular Latin square design, the computation is as follows:

Correction Factor	$G^2/np^2$	[2]
Total SS	$\sum x^2$	[3]

<sup>12</sup>Hays, Statistics for Psychologists, pp. 356-375.



Order	$\sum A^2/np$	[4]
Treatment	$\sum B^2/np$	[5]
Cell Total	$\sum (AB_{ij})^2/n$	[6]
Paired Observations	$\sum P_{m(k)}^2/p$	[7]
Row	$\sum C_k^2/np$	[8]
$SS_a = [4] - [2]$		[9]
$SS_b = [5] - [2]$		[10]
$SS_{groups} = [8] - [2]$		[11]
$SS_{subj. w.groups} = [7] - [8]$		[12]
$SS_{error(within)} = [3] - [6]$		[13]

The summations in all cases are over the whole experiment.

$P_{m(k)}$  represents the sum of the  $q$  observations on person  $m$  in group  $k$ .  $G$  is the grand total;  $A$ ,  $B$ ,  $C$ , represent the sum of the  $np$  observations for their corresponding sources of variation.

The ratio for the  $F$  test is computed for each source of variation in the usual manner by taking the ratio of two mean square values of interest.  $F(u,v)$  will designate the  $F$  value where  $u$  is the number of degrees of freedom in the numerator of the ratio and  $v$  the number of degrees of freedom in the denominator.  $F(1,16)$  then for this experiment is:

$$F(1,16) = MS_a / MS_{error(within)}$$

$$F(1,16) = MS_b / MS_{error(within)}$$

$$F(1,16) = MS_{groups} / MS_{subj w.groups}$$



In order to test the error term for homogeneity of variance, [12] is also computed so as to provide the parts that may be checked by an  $F_{\max}$  test.<sup>13</sup> This is accomplished by summing only over Group I and over Group II rather than over the whole experiment. In this way the two components so derived for [13] are tested at  $F(8,8)$  for an appropriate significance level. The magnitude of this  $F$  ratio indicates the extent to which the observed data conform to the model, and thus provides a partial check on the appropriateness of the model.

Since the values of the decision month are positive integers, a skewed distribution was suspected and found to be present. The nature of the decision is also such that higher variances would be expected to be associated with higher means. The values were, therefore, transformed following a standard procedure by taking the square root, to make the variances more homogeneous and at the same time to normalize the distribution.<sup>14</sup> It is this square root value which is used in the analysis of variance computations, and reported in the tables.

It is of interest to compare the rank order of the inventory control policies at the decision month and the rank order at the twenty-fourth month for each individual. This comparison gives a measure of each individual's

---

<sup>13</sup>Winer, Statistical Principles, pp. 92-96, 542.

<sup>14</sup>Ibid., pp. 218-222.





consistency between the ranking decision made at the decision month and that ranking decision made when full information was available. These two rankings for each participant are compared by means of a Spearman rank correlation coefficient test.<sup>15</sup> The computational procedure is as follows:

$$r_s = 1 - \left( \frac{6 \sum D_i^2}{N (N^2 - 1)} \right) \quad [14]$$

where  $D_i$  is the difference in ranking for the  $i^{\text{th}}$  inventory control policy, and  $N = 12$ , the number of inventory control policies observed. Once computed, these values for  $r_s$  are treated as data for analysis in the Latin square design previously discussed.

Since the  $r_s$  values are known to have a skewed distribution, a transformation is made on each of the  $r_s$  values in order to place the data on symmetric scale, so as to normalize the distribution.<sup>16</sup> It is these transformed values for  $r_s$  which are used in the analysis of variance computations. The transformation is:

$$r_s \text{ transformed} = .5 \left( \ln \frac{1+r_s}{1-r_s} \right) \quad [15]$$

This transformation is a routine procedure used on correlation coefficients and is justified and discussed in

---

<sup>15</sup>Hays, Statistics for Psychologists, pp. 643-646.

<sup>16</sup>R. G. D. Steel and J. H. Torrie, Principles and Procedures of Statistics (New York: McGraw-Hill Book Company, 1960), pp. 188-189.





detail in the reference.

In order to find the extent to which the members of each group ranked the policies the same way under the same conditions, Kendall's W coefficient of concordance<sup>17</sup> was computed for each cell of the Latin square. With this single statistic computed for each of the four cells, some comments can be made concerning the homogeneity of decisions under the conditions of each cell. The computation for Kendall's W statistic is as follows:

$$W = \left( \frac{12 \sum_j T_j^2}{m^2 N(N^2 - 1)} \right) - \frac{3(N+1)}{N-1} \quad [16]$$

where for this experiment  $m = 9$ , the number of individuals doing the ranking;  $N = 12$ , the number of policies ranked; and  $T_j$  is the sum of the nine individual rankings given the  $j^{\text{th}}$  policy (i.e., the sum of column  $j$  in Tables 9 through 16). The statistic  $W$  is based on the extent of variability among the respective sums of the ranks. The basic relationship is:

$$W = \frac{\text{variance of rank sums}}{\text{maximum possible variance of rank sums}}$$

which reduces to [16]. From this it can be seen that  $W$  indicates the degree of concordance, or how well the group members agreed among themselves as to the rankings.

Consider another statistic. For each of the

---

<sup>17</sup>Hays, Statistics for Psychologists, pp. 656-658.



$\binom{9}{2} = 36$  pairs of subjects in a cell, calculate the Spearman coefficient for their two policy rankings at the decision month. The average of these coefficients is a measure of the agreement among subjects. This average may be computed directly from Kendall's W statistic by:

$$\text{average } r_s = \frac{mW-1}{m-1}$$

In the interpretation of the result of this computation it should be noted that this experimental group was not attempting to apply a common criterion of excellence in making their decisions. Individual best judgment and experience in the decision process guided the problem solution. The course which the subjects attended neither taught nor encouraged standards in decision making or in performance goals.

The difference between the two values of  $r_s$  just discussed is that the single value of average  $r_s$  derived by way of Kendall's W statistic is a measure of concordance among all nine participants within a cell, whereas the individual values of  $r_s$  derived from Spearman's test are a measure of each person's consistency.



## VI. RESULTS

The data analyzed and the results of the analysis appear in the following tables. The results were computed by PL/I routines, using single precision floating point arithmetic. Where the tables appear in the form of the Latin square, each cell is represented as in Figure 12. The order of the participants in the listings is constant over the tables presented; that is, the  $i^{\text{th}}$  data element always refers to the  $i^{\text{th}}$  individual regardless of which table is referenced. The within-cell mean and standard deviation, corrected for sample size, are included in the data lists.

In the analysis of variance we are examining the two treatment hypotheses:

$$H_0 : \mu_P = \mu_D$$

$$H_1 : \mu_P > \mu_D$$

That is, the null hypothesis is that the mean associated with the display unit is equal to the mean associated with the printer output. Since we believe that the display unit is, in fact, the better decision-making tool of the two, the alternative hypothesis is that the mean associated with the printer output is strictly greater than the mean associated with the display unit. We are also comment-





ing on the two order hypotheses:

$$H_0 : \mu_{01} = \mu_{02}$$

$$H_1 : \mu_{01} > \mu_{02}$$

involving the means associated with Order 1 and Order 2. The usual method in statistical hypotheses testing involves setting the significance level of the test in advance of obtaining the data. The convention used in the analysis of variance deviates somewhat from this formality. The value of F is reported to exceed a specified percentile by a comparison with tables of critical values. The choice of the significance level is in this way determined partly by the data itself.<sup>18</sup> This allows each reader to establish his own significance level and to judge the results thereby. We will consider a conclusion to be more surely established if the probability of its truth is higher. This methodology does not allow a measure of the power of the test; however, the procedure is valid for estimating the probabilities of the observation in relationship to the assumed sampling distribution. In this preliminary investigation we would be encouraged to find significance at the .05 level, and the significance levels actually attained are noted in the tables.

In this analysis of variance, we are principally interested in investigating the effect of the treatment on the subjects. We will also be interested in commenting on

---

<sup>18</sup>Winer, Statistical Principles, p. 55.





the effect of order on the subjects. The results which correspond to these effects will be referred to as the within-subject sources of variation. We have also reported the effect which corresponds to differences between Group I and Group II as the between-subject source of variation. The results of the various analysis of variance computations are shown in Tables 2, 4, 6, and 8.

In the analysis of variance for the ordinal value of the decision month, we were compelled to use a transformation to produce a more symmetric distribution of values, as noted in Chapter V. Both the original values for this month and the transformed values are presented in Tables 1a and 1b, respectively. These values represent how many months of data the participants needed prior to committing themselves to a decision on the rank ordering that they would recommend. The question is: How early in simulated time can a decision be made?

The null hypothesis that the variance is equal over the groups cannot be rejected, as the  $F_{\max}$  test indicates (Table 2). The significance level of .10 appears to be marginal, and we cannot really say that we accept the null hypothesis either. However, we are sufficiently satisfied with this result to have confidence in the validity of the analysis of variance for the treatment factor. For the treatment factor,  $H_0$  is not supported at the .01 level. We can say with a high degree of assurance that the mean decision month arrived at by using the display unit was



TABLE 1a  
ORDINAL VALUE OF DECISION MONTH

	Order 1		Order 2
<u>Group I</u>	13		14
	6		14
	8		13
	10	Display	11
	12	mean=10.222	17
	10	$\sigma = 4.086$	14
	19		14
	8		12
	6		10
<u>Group II</u>	9		8
	9		7
	10		9
	8	Printer	8
	13	mean= 9.333	10
	8	$\sigma = 2.121$	7
	6		7
	12		7
	9		10

indeed less than that using the printer output. From another viewpoint, the participants on the average needed to look at less data volume with the display unit to make a committing decision than they did with the printer output. There is only a 1 percent probability that these observed differences are due to chance effects. No effect can be attributed to order, as indicated by the significance level of .25 for order.

Reference to Table 1a gives some additional information on the characteristics of the data within each cell. Within each group, the mean decision month is lower for the presentation on the display unit. The standard deviation



TABLE 1b  
TRANSFORMED ORDINAL VALUE OF DECISION MONTH

	Order 1		Order 2	
<u>Group I</u>	3.605		3.741	
	2.449		3.741	
	2.828		3.605	
	3.162	Display	3.316	Printer
	3.464	mean=3.145	4.123	mean=3.626
	3.162	$\sigma = .607$	3.741	$\sigma = .282$
	4.358		3.741	
	2.828		3.464	
	2.449		3.162	
<u>Group II</u>	3.000		2.828	
	3.000		2.645	
	3.162		3.000	
	2.828	Printer	2.828	Display
	3.605	mean=3.037	3.162	mean=2.840
	2.828	$\sigma = .346$	2.645	$\sigma = .219$
	2.449		2.645	
	3.464		2.645	
	3.000		3.162	

does not follow this pattern however. For Group I, the standard deviation,  $\sigma$ , of values for the display unit presentation, 4.086, is considerably higher than for the printer output, 2.048, while just the opposite is true for Group II (1.269 vs. 2.121). In other words, both groups had a higher  $\sigma$  for Order 1 than they had for Order 2. Apparently there was some maturation between Order 1 and Order 2 which was reflected in  $\sigma$ , but not in the mean in both groups. The participants who were presented with Order 1 on the display unit were encountering both a unique problem and a unique way of analyzing the data, whereas Group II presented with Order 1 used a familiar data



TABLE 2  
ANALYSIS OF TRANSFORMED DECISION MONTH

Variation Source	SS	df	MS	F(1,16)	Signifi- cant at
<u>Between Subjects</u>					
Groups	1.797	1	1.797	8.485	.05
Subjects w/in Groups	3.390	16	0.211		
<u>Within Subjects</u>					
Order	0.181	1	0.181	1.877	.25
Treatment	1.034	1	1.034	10.709	.01
Error (Within)	1.546	16	0.096		

Homogeneity of Variance,  $F_{\max}$  Test:

Variation Source	SS	df	F(8,8)	Signifi- cant at
<u>Within Subjects</u>				
Error (Within) Group I	1.148	8		
Error (Within) Group II	.398	8	2.884	.10
	1.546	16		





presentation for data analysis. Note that one very cautious participant, who waited until the nineteenth month to commit himself to a ranking, did quite a bit of damage to the standard deviation for the Group I, Order 1 cell.

The elapsed clock time which was required for a participant to commit himself to a ranking is given in Table 3, and its analysis of variance in Table 4. The null hypothesis of homogeneity of variance cannot be rejected by the  $F_{\max}$  test, allowing us to accept the results of the analysis of variance.

The treatment factor is shown to be significant at much beyond the .001 level. The mean time to decision using the display unit is with great certainty less than that using the printer output. This indicates that the amount of time spent in making a decision was significantly reduced in this experiment by using the display device, which may point to an ability to assimilate a large quantity of data and to correlate these data by using display techniques, as opposed to printer output. The convenience of having virtually instantaneous recall of data displays by using the Programmed Function Keyboard is certainly a consideration in the interpretation of the results. Pushing buttons is just inherently faster than paging through a book of data, however well arranged and indexed the book may be. However much or little this consideration entered into the results, we still can give credence to the



TABLE 3  
ELAPSED TIME IN MINUTES TO DECISION MONTH

	Order 1		Order 2	
<u>Group I</u>	63		67	
	35		85	
	75		127	
	68	Display	72	Printer
	88	mean=64.666	115	mean=89.777
	79	$\sigma=18.500$	70	$\sigma=23.156$
	81		114	
	52		89	
	41		69	
<u>Group II</u>	52		41	
	69		29	
	60		47	
	66	Printer	31	Display
	109	mean=73.444	43	mean=39.222
	61	$\sigma=19.255$	48	$\sigma=6.457$
	95		38	
	61		37	
	88		39	

proposition that data can be correlated faster and retained better from a properly programmed display unit.

Unsolicited comment from individual participants supported this conjecture without exception. The observation of the author is that the dynamic graphs gave the participant a much better intuitive feel for the situation, and that he was more likely to retain this impression and not have to refer to past data repeatedly. One consideration in this observation was the program control which always started the dynamic graphs at month 1 and brought them up to the current month in increments of one second. This forced a continual review of the history and



TABLE 4  
ANALYSIS OF DECISION TIME

Variation Source	SS	df	MS	F(1,16)	Signifi- cant at
<u>Between Subjects</u>					
Groups	3927.125	1	3927.125	9.161	.01
Subjects w/in Groups	6858.125	16	428.632		
<u>Within Subjects</u>					
Order	186.750	1	186.750	0.861	not
Treatment	7921.000	1	7921.000	36.531	<<.001
Error (Within)	3469.250	16	216.828		

Homogeneity of Variance,  $F_{\max}$  Test:

Variation Source	SS	df	F(8,8)	Signifi- cant at
<u>Within Subjects</u>				
Error (Within) Group I	1806.43	8		
Error (Within) Group II	1662.78	8	1.086	not
	3469.21	16		





derivative of the measure under consideration and undoubtedly reinforced past impressions. It was seldom that a participant asked that the graph be stopped at a month prior to his current month so that he could review the static situation as of that past instant. As a matter of interest, it was noted when the experiment was well under way that the more experienced participants used the dynamic graphs extensively, where the less experienced participants relied on the tabular listings presented on the display unit.

Referring to Table 3, we note that both the mean time to make this decision and the standard deviation within the cell are considerably reduced in favor of the display unit over the printer output. Group I, in going from display unit presentation to printer output presentation, showed a mean increase of about twenty-five minutes to make a decision, and an increase in  $\sigma$ . Group II, in going from printer output to display unit presentation, showed a mean decrease of about thirty-four minutes and a very sharp drop in  $\sigma$ . It might be argued that in spite of the maturation between Order 1 and Order 2, Group I did better on the display unit, and that the very noticeable difference between Order 1 and Order 2 statistics for Group II is due only in part to this maturation. It is interesting to note that only one participant (Participant F) made a faster decision using the printer output than using the display unit. If we examine the data in terms





of order, it can be seen that within each order the display device gave more favorable results in terms of mean and standard deviation than did the printer output.

The statistics of Tables 5 and 6 for the total elapsed time to complete the problem after the decision month, and its analysis of variance, are included for interest. It can reasonably be assumed that there was some speed-up by participants after they had made their decision to commit, simply to get to the end to see how well they did. There is some evidence of this speed-up in the time data. How much of this is due to increased familiarity with the problem at hand and how much is due to impatience to get to the final result is difficult to say. The elapsed time after the decision month was tempered by the requirement that the data be ranked at each month. From personal observation, the participants appeared to be conscientious about following the spirit and the letter of the instructions, but relieved that the big decision had been reached, and were in a hurry to finish the twenty-four months to check their final ranking against their decision month ranking.

Table 6 shows that the hypothesis of homogeneous variance cannot be rejected, and we can go directly to comment on the within-subject sources of variation. Here the order statistics show that  $H_0$  cannot be supported at the .01 level, which is sufficient to support a statement that we accept  $H_1$ . The order factor is thus seen to show



TABLE 5  
TIME IN MINUTES TO COMPLETE THE PROBLEM  
AFTER THE DECISION MONTH

	Order 1	Order 2
<u>Group I</u>	62	41
	71	47
	87	54
	64 Display	55 Printer
	50 mean=59.777	28 mean=44.555
	64 $\sigma=17.788$	32 $\sigma=13.314$
	21	70
	56	35
	63	39
<u>Group II</u>	58	57
	36	51
	83	55
	65 Printer	52 Display
	68 mean=72.666	54 mean=51.000
	106 $\sigma=22.237$	46 $\sigma=11.135$
	93	66
	54	53
	91	25

significance here, which is reasonable and consistent with our previous comments on the effects of order. In this test the treatment factor was completely without significance which is also a reasonable result in view of the observation concerning the impatience to finish the problem. Reference to Table 5 shows a completely order-dependent data structure.

In Table 8, the analysis of variance statistics for the transformed Spearman rank correlation coefficient,  $r_s$ , are presented. The  $F_{\max}$  test is satisfactory in our investigation of equal variances, so we turn our attention



TABLE 6  
ANALYSIS OF TIME TO COMPLETE PROBLEM  
AFTER DECISION MONTH

Variation Source	SS	df	MS	F(1,16)	Signifi- cant at
<u>Between Subjects</u>					
Groups	841.000	1	841.000	3.882	.10
Subjects w/in Groups	3466.000	16	216.625		
<u>Within Subjects</u>					
Order	3061.750	1	3061.750	9.018	.01
Treatment	93.437	1	93.437	0.275	not
Error (Within)	5431.812	16	339.488		

Homogeneity of Variance,  $F_{\max}$  Test:

Variation Source	SS	df	F(8,8)	Signifi- cant at
<u>Within Subjects</u>				
Error (Within) Group I	2513.78	8		
Error (Within) Group II	2918.00	8	1.160	not
	5431.78	16		





to the within-subject sources of variation. Here our hypotheses must be modified to read:

Treatment Factor

$$H_0: \mu_D = \mu_P$$

$$H_1: \mu_D > \mu_P$$

Order Factor

$$H_0: \mu_{01} = \mu_{02}$$

$$H_1: \mu_{02} > \mu_{01}$$

The hypothesis of equal means for the treatment factor,  $H_0$ , cannot be supported at the .05 level. With a significance level of .05, we obviously do not have the clear mandate that our other treatment factor statistics have given, but we are willing to claim that the evidence points towards rejection of  $H_0$  and acceptance of  $H_1$ ; that is the mean correlation coefficient is higher using the display unit than using the printer output. The values of  $r_s$  give a correspondence between the participant's ranking of the policies at the decision month and his ranking at the last month of the simulation data, month 24. This, then, is a measure of the consistency between these two rankings. It is also a measure of the participants discrimination ability - that is his ability to decide whether he has enough information to commit himself to a ranking or not. A decision to commit too soon in relation to each individual's ability and ranking criteria would, in most cases, result in a poor correlation coefficient, whereas being overly cautious and waiting beyond the point where he had sufficient information could not be expected to materially improve the correlation coefficient. Thus, we might say that one interpretation of a low  $r_s$  would be that the





participant committed himself too early. Other interpretations are, of course, that he simply used poor judgment in his ranking, or that he materially changed his ranking criteria after the decision month. Participants were cautioned to use a consistent ranking schema throughout. As an extreme example, it was pointed out to the class that to rank the policies based only on lost sales data through the decision month, then to abandon that schema and to rank the policies only on number of purchase orders generated would not be showing responsible judgment. On the basis of these comments, we should be able to narrow our consideration of a principal cause of low  $r_s$  to either too early commitment or poor ranking judgment at some point. Both of these essentially are measures of decision-making ability and we can accept either one or both as reasonable interpretations of a low correlation coefficient.

From Table 8, we can see that the order factor is highly significant; the probability that the mean of Order 1 is equal to the mean of Order 2 is something less than 0.1 percent. Thus we accept the alternative hypothesis,  $H_1$ , that the mean of Order 2 is greater than the mean of Order 1. The maturation between Order 1 and Order 2 would appear to play a large role in the decision making consistency that is being measured. That is, the ability to meaningfully rank a set of policies grows with practice, and moreover, improves more than the actual time and simulation decision month considerations previously



TABLE 7a  
SPEARMAN RANK CORRELATION COEFFICIENT  $r_s$

	Order 1		Order 2
<u>Group I</u>	.615		.755
	.405		.832
	.118		.685
	.412	Display	.559
	.447	mean=.501	.965
	.671	$\sigma$ =.195	.706
	.713		.552
	.734		.055
	.475		.713
<u>Group II</u>	-.062		.678
	.398		.741
	-.580		.811
	.279	Printer	.720
	.209	mean=.126	.720
	.188	$\sigma$ =.302	.923
	.405		.762
	.209		.783
	.090		.706

measured. Reference to Table 7a shows that the mean for both groups increased from Order 1 to Order 2, but that the standard deviation for both groups under the display unit treatment was less than the standard deviation for both groups under the printer output presentation. The data for the standard deviation in this table are of the same general pattern as the data for the standard deviation in Table 3, but Table 7a is the only table examined so far in which the means within both groups increased with order. One prominent difference in this comparison is that Group II performed very poorly under the printer output for Order 1 and quite well under the display output for



TABLE 7b

TRANSFORMED SPEARMAN RANK CORRELATION COEFFICIENT  $r_s$ 

	Order 1		Order 2	
<u>Group I</u>	0.717		0.985	
	0.430		1.195	
	0.119		0.839	
	0.438	Display	0.632	Printer
	0.481	mean=.594	2.014	mean=.901
	0.813	$\sigma=.266$	0.879	$\sigma=.524$
	0.893		0.621	
	0.937		0.056	
	0.517		0.893	
<u>Group II</u>	-0.063		0.825	
	0.421		0.953	
	-0.663		1.130	
	0.287	Printer	0.908	Display
	0.212	mean=.124	0.908	mean=1.030
	0.191	$\sigma=.332$	1.609	$\sigma=.236$
	0.430		1.001	
	0.212		1.053	
	0.091		0.879	

Order 2. The lowest and highest means and standard deviations for all four cells are found in the Group II data. The use of printer output for Group II had the highest standard deviation, and the lowest mean, while the use of the display device for this group had the highest mean and the lowest standard deviation. Interpretation of the data from Table 7a in conjunction with the treatment factor data from Table 8 gives more credibility to the acceptance of  $H_1$  for the treatment factor.

Tables 9 through 12 show the rankings of each policy by the participants at the decision month. Each of the tables represents one cell of the Latin square. For



TABLE 8

ANALYSIS OF TRANSFORMED SPEARMAN RANK  
CORRELATION COEFFICIENT,  $r_s$

Variation Source	SS	df	MS	F(1,16)	Signifi- cant at
<u>Between Subjects</u>					
Groups	0.262	1	0.262	2.759	.25
Subjects w/in Groups	1.522	16	0.095		
<u>Within Subjects</u>					
Order	3.310	1	3.310	20.546	<.001
Treatment	0.804	1	0.804	4.993	.05
Error (Within)	2.577	16	0.161		

Homogeneity of Variance,  $F_{\max}$  Test:

Variation Source	SS	df	F(8,8)	Signifi- cant at
<u>Within Subjects</u>				
Error (Within) Group I	1.854	8		
Error (Within) Group II	.723	8	2.563	.25
	2.577	16		





example, in Table 9, Participant A ranked policy 1 as eleventh and policy 2 as seventh at the decision month. To demonstrate the correspondence among the tables, reference to Table 1a and Table 3 shows that Participant A reached this ranking at month 13 after an elapsed time of sixty-three minutes. For Tables 9 through 12, Kendall's W statistic and more importantly, the derived value of average  $r_s$  is indicated. The values of average  $r_s$  from these tables are summarized as follows in our usual format:

	Order 1	Order 2
Group I	Display $r_s = .574$	Printer $r_s = .664$
Group II	Printer $r_s = .543$	Display $r_s = .829$

As was the case in Table 7a, the effect of order is evident here from the increase in average  $r_s$  from Order 1 to Order 2 in both groups, and the most prominent increase is in the movement from printer to display unit output in Group II. Within the consideration of order, the cell which represents the display unit data had the higher value of  $r_s$  for both Order 1 and Order 2. This would tend to indicate for this experiment that whatever the order, a more homogeneous ranking was made by the participants under display unit presentation than under printer output presentation.



TABLE 9

DATA AND RESULTS FROM KENDALL'S W STATISTIC  
 COMPUTATIONS FOR UPPER LEFT QUADRANT  
 OF LATIN SQUARE; GROUP I,  
 ORDER 1, DISPLAY UNIT

Partici- pant	Policy Number											
	1	2	3	4	5	6	7	8	9	10	11	12
A	11	7	6	2	10	4	9	12	1	8	5	3
B	6	2	7	4	10	5	12	8	9	11	1	3
C	7	3	8	1	9	6	12	11	2	10	4	5
D	8	4	9	1	10	5	11	12	3	7	6	2
E	12	10	2	3	6	7	5	11	4	8	9	1
F	3	1	8	5	9	6	12	11	7	10	2	4
G	7	5	8	11	6	1	12	9	3	10	2	4
H	9	2	8	1	7	6	12	11	3	10	5	4
I	9	2	6	4	10	7	12	11	5	8	1	3

Summary of ranking by participants at decision month;

$$W = .621$$

$$\text{average } r_s = .574$$



TABLE 10

DATA AND RESULTS FROM KENDALL'S W STATISTIC  
 COMPUTATIONS FOR UPPER RIGHT QUADRANT  
 OF LATIN SQUARE; GROUP I, ORDER 2,  
 PRINTER OUTPUT

Partici- pant	Policy Number											
	1	2	3	4	5	6	7	8	9	10	11	12
A	9	5	12	7	11	10	1	3	4	8	6	2
B	6	2	12	8	9	7	1	3	5	10	11	4
C	7	4	11	8	10	9	1	3	5	6	12	2
D	6	7	5	8	11	10	2	3	4	9	12	1
E	6	1	12	9	7	10	3	2	5	11	8	4
F	3	2	12	9	8	5	4	1	7	10	11	6
G	6	3	12	5	8	10	2	1	7	11	9	4
H	10	9	6	11	8	3	2	1	7	5	12	4
I	7	3	6	8	11	10	1	4	5	9	12	2

Summary of ranking by participants at decision month;

$$W = .701$$

$$\text{average } r_s = .664$$



TABLE 11

DATA AND RESULTS FROM KENDALL'S W STATISTIC  
 COMPUTATIONS FOR LOWER LEFT QUADRANT  
 OF LATIN SQUARE; GROUP II, ORDER 1,  
 PRINTER OUTPUT

Partici- pant	Policy Number											
	1	2	3	4	5	6	7	8	9	10	11	12
J	9	5	6	1	7	3	11	12	4	10	8	2
K	6	1	9	2	10	3	11	12	5	8	4	7
L	7	3	2	12	1	4	6	10	8	5	9	11
M	4	1	6	5	7	8	12	10	9	11	3	2
N	6	1	7	3	10	4	11	12	5	8	9	2
O	4	1	6	3	9	7	12	10	8	11	2	5
P	7	4	9	3	10	6	12	11	5	8	1	2
Q	4	1	7	2	8	5	11	12	9	10	6	3
R	7	2	9	1	10	6	11	12	3	8	4	5

Summary of ranking by participants at decision month;

$$W = .593$$

$$\text{average } r_s = .543$$





TABLE 12

DATA AND RESULTS FROM KENDALL'S W STATISTIC  
COMPUTATIONS FOR LOWER RIGHT QUADRANT  
OF LATIN SQUARE; GROUP II,  
ORDER 2, DISPLAY UNIT

Partici- pant	Policy Number											
	1	2	3	4	5	6	7	8	9	10	11	12
J	9	3	6	8	10	11	2	4	5	7	12	1
K	7	3	6	9	11	12	2	4	5	8	10	1
L	6	2	7	8	11	10	4	1	5	9	12	3
M	6	1	3	9	11	10	4	2	5	8	12	7
N	6	4	7	8	11	9	1	3	5	10	12	2
O	3	2	6	7	9	11	5	1	4	12	10	8
P	6	1	8	9	12	11	2	4	5	7	10	3
Q	5	2	6	8	9	11	7	1	3	10	12	4
R	6	3	7	9	11	10	2	4	5	8	12	1

Summary of ranking by participants at decision month;

$$W = .848$$

$$\text{average } r_s = .829$$



It should be emphasized that the value of average  $r_s$  derived by way of Kendall's W statistic (Tables 9 through 12) is a measure of concordance among all nine participants within a cell, whereas the values of  $r_s$  derived from Spearman's test (Table 7a) are a measure of each person's consistency. It is also worthwhile commenting again that the concordance was not measured against predetermined criteria, but rather reflects the degree to which experienced decision makers tended to agree under different treatments but using their own best judgment.

Tables 13 through 16 give the results of the decision process at the end of the twenty-fourth month when the decider had full information available to him. These tables represent the data in the same manner as did Tables 9 through 12. That is, in Table 13, Participant E ranked policy 1 as last, policy 2 as tenth and policy 3 as ninth. The resulting average  $r_s$  is represented in our usual Latin square format as:

	Order 1	Order 2
Group I	Display .639	Printer .856
Group II	Printer .485	Display .947

Here the pattern is much the same as it was for the decision month average  $r_s$  (see page 72 ), except that the



TABLE 13

DATA AND RESULTS FROM KENDALL'S W STATISTIC  
 COMPUTATIONS FOR UPPER LEFT QUADRANT  
 OF LATIN SQUARE AT TWENTY-FOURTH  
 MONTH; GROUP I, ORDER 1  
 DISPLAY UNIT

Partici- pant	Policy Number											
	1	2	3	4	5	6	7	8	9	10	11	12
A	12	6	7	9	10	8	5	11	1	4	2	3
B	8	1	11	12	7	5	9	10	3	6	2	4
C	11	5	10	12	6	9	7	8	1	3	2	4
D	12	5	9	11	7	4	8	10	2	6	1	3
E	12	10	9	4	7	11	5	6	3	8	1	2
F	6	1	8	12	10	4	11	9	3	7	2	5
G	7	4	12	10	5	8	11	6	2	9	1	3
H	12	5	11	4	6	8	10	9	2	7	1	3
I	9	1	11	12	7	6	8	10	2	4	3	5

Summary of ranking by participants at twenty-fourth  
 month:

$$W = .679$$

$$\text{average } r_s = .639$$



TABLE 14

DATA AND RESULTS FROM KENDALL'S W STATISTIC  
 COMPUTATIONS FOR UPPER RIGHT QUADRANT  
 OF LATIN SQUARE AT TWENTY-FOURTH  
 MONTH; GROUP I, ORDER 2  
 PRINTER OUTPUT

Partici- pant	Policy Number											
	1	2	3	4	5	6	7	8	9	10	11	12
A	7	2	11	8	6	12	3	1	5	10	9	4
B	6	2	10	9	7	11	3	1	4	12	8	5
C	7	2	10	9	6	11	3	1	5	12	8	4
D	7	1	10	9	6	11	3	2	5	12	8	4
E	6	2	10	9	7	11	3	1	4	12	8	5
F	4	2	8	10	6	11	5	1	3	12	9	7
G	7	2	6	10	12	9	3	5	4	8	11	1
H	7	2	9	6	4	12	5	1	3	11	10	8
I	6	1	10	9	7	11	3	2	4	12	8	5

Summary of ranking by participants at twenty-fourth  
 month:

$$W = .872$$

$$\text{average } r_s = .856$$





TABLE 15

DATA AND RESULTS FROM KENDALL'S W STATISTIC  
 COMPUTATIONS FOR LOWER LEFT QUADRANT  
 OF LATIN SQUARE AT TWENTY-FOURTH  
 MONTH; GROUP II, ORDER 1  
 PRINTER OUTPUT

Partici- pant	Policy Number											
	1	2	3	4	5	6	7	8	9	10	11	12
J	12	8	11	5	9	10	4	3	6	7	1	2
K	10	1	12	11	7	6	8	9	2	4	3	5
L	12	11	9	5	8	10	1	4	3	7	6	2
M	6	5	3	12	11	10	9	1	7	8	2	4
N	12	2	11	10	6	7	8	9	3	5	1	4
O	8	2	9	12	11	10	7	4	3	6	1	5
P	9	1	11	12	7	5	8	10	2	4	3	6
Q	8	1	11	12	7	5	9	6	3	10	2	4
R	12	8	11	10	5	9	6	7	2	4	1	3

Summary of ranking by participants at twenty-fourth month:

$$W = .542$$

$$\text{average } r_s = .485$$



TABLE 16

DATA AND RESULTS FROM KENDALL'S W STATISTIC  
 COMPUTATIONS FOR LOWER RIGHT QUADRANT  
 OF LATIN SQUARE AT TWENTY-FOURTH  
 MONTH; GROUP II, ORDER 2  
 DISPLAY UNIT

Partici- pant	Policy Number											
	1	2	3	4	5	6	7	8	9	10	11	12
J	7	4	10	9	6	11	2	1	5	12	8	3
K	6	2	10	9	7	11	3	1	5	12	8	4
L	6	2	10	9	7	11	3	1	5	12	8	4
M	5	1	9	8	7	11	4	2	3	12	10	6
N	7	2	11	9	6	10	3	1	5	12	8	4
O	4	2	9	8	6	11	5	1	3	12	10	7
P	6	1	10	9	7	11	3	2	4	12	8	5
Q	5	2	11	8	7	9	3	1	4	12	10	6
R	6	1	10	9	7	11	3	2	4	12	8	5

Summary of ranking by participants at twenty-fourth  
 month:

$$W = .953$$

$$\text{average } r_s = .947$$



effect of Group II going from printer output at Order 1 to display unit output at Order 2 is much more pronounced. There was a moderate increase in the average  $r_s$  for Group I going from display to printer output, which may be ascribed in part to maturation. However, the average  $r_s$  almost doubling in Group II when going from printer output to display output may be more than can plausibly be ascribed to maturation alone. In comparing these results with the average  $r_s$  results from Tables 9 through 12, it would be well to remember that the rankings from those tables did not represent the rankings made all at the same month. In the case of the rankings at the twenty-fourth month, the values of average  $r_s$  may be a more valid measure of concordance since the decision relating to the ranking is not confused with the additional decision of when to indicate a decision month. In the rankings at the twenty-fourth month, the deciders all had the same amount of information available to them, which was not the case for the decision month rankings.



point to the use of a display presentation when economy of time or simply volume of printed output is a serious constraint on the system or the decision maker.

While the two results reported above have the rigorous respectability of high statistical significance, the next results to be discussed are at least as important in the evaluation of the experiment. These are the results which answer the question of whether a better decision can be made with a display device. Because of the criteria used, we will claim that a decision at the month of commitment that is more consistent with the final decision is a better decision, and also that a decision which exhibits evidence of higher concordance among group members is a better decision. However, the two major results listed above would be convincing in their own right even if the decisions turned out to be no worse using a display device.

The results from Kendall's W statistic and from Spearman's rank correlation statistic show that individuals tended to make more consistent decisions with themselves and with their group using display device presentation techniques. Casual inspection indicates the conclusions established by the statistical analysis; within order, individuals made more consistent decisions using the display. It is equally clear that within group the increase in correlation in going from printer output to display output (Group II) far exceeds the increase in correlation in going from display output to printer output





(Group I) for both of these statistics.

The evidence from the Spearman and the Kendall tests definitely indicate better decision making under display device presentation. With this result we conclude the evaluation of the statistical results.

The economics of a system of display devices for decision making will not be explored here, however, it is evident that the very specialized research equipment used for this experiment is both expensive and unnecessarily elaborate for an operational system. The prices and characteristics of display equipment that will serve the purpose very well are frequently published.<sup>19</sup>

The minimum display unit for implementation of an information system of this general type should have alphanumeric display capabilities and a programmed function keyboard, or equivalent means of easy display selection. The size of the display face is crucial to the extent that it must be able to contain enough material to be of interest and still allow character size and spacing to enhance readability. For instance, the IBM 2250 used in this experiment has a display face twelve inches square with a maximum capability of fifty-two lines of seventy-four characters each. The information in the displays (Figures 6, 7, 8, 9) is rather densely packed and is digestable only by someone sitting in the console chair immediately

---

<sup>19</sup>Adams Associates, Computer Characteristics Quarterly, Second Quarter 1968 (Watertown, Massachusetts: Keydata and Adams Associates, Inc., 1968), pp. 179-187.



in front of the display face. A smaller display face would mean that displays would have to be segmented; the same information in smaller characters on a smaller display face would be the wrong compromise. With segmented displays, more programmed function keys would be needed and in some ways the problem of how to ask for a particular display becomes more complicated for the user. It is unfortunate that the great majority of the pure alphanumeric display units have small display faces - eight inches by six inches appears to be a popular size. Other features of the IBM 2250, such as an alphanumeric type-writer keyboard, line-drawing capability, and a light pen, would serve only to detract from the simplicity and utility of the system and add only to the cost of this application.

In addition to the display device proper, this experiment used other system facilities. The display unit had a self-contained buffer of 8,192 characters. Of this, a maximum of 2,000 characters of buffer storage were used at any one time. The display program in the main storage of the IBM 360 Model 40 used approximately 13,000 characters for program and 21,000 characters for tables. An additional 46,000 characters of disk storage were used for table overlays.

How generalizable are the results of this experiment? The display system evidently achieved the objective of presenting a complex situation, which involved many inter-relationships, in a manner such that the key concepts and



fundamental correlations were clearly understandable. The display system appeared to facilitate interpretation and extrapolation of the relevant data; a significant element in the decision process. The reduction of reaction time of top-level decision makers in this environment is both an interesting result and an important objective of any executive display system. While there is an obligation to report these results in relation to the experiment group itself, the composition of this group allows us to extend the results to a wider group of individuals. It would appear reasonable to apply the results to that class of individuals who are experienced in and actively making decisions in the inventory control field. It is tempting to extend these results to decision making in general; however, that extension will have to await further investigation and experimentation.

What direction should further work take? One point of great interest would be the exploration of the differential cost or savings of decisions using display units and printer output. This is a rather difficult area to define in that dollar values and weightings must be assigned not only to such elements as a reduction in inventory valuation and the cost of lost sales, but also to the generated purchase actions, availability of material, timeliness of decision, system cost and other influencing factors. The advantage of having such a differential cost at hand is obvious; just how to go about obtaining it is not.





Statistics on the frequency of use of the various displays should be collected, both automatically and by experimenter observation. The correlation of the frequency of use by display type with the individual's consistency of decision would be most important for the design of extensions of this system. Unfortunately, the importance of this statistic was not realized until this experiment was well under way.

Whatever the extension of the program, there should be the capability for the decider to request a hard copy of any display he wishes. If line drawing capability is used, this, of course, implies the availability of the equivalent of a line plotter for hard copy output. This requirement is more operational than experimental. We have no doubt as to the utility of such a feature for the decider in an operational environment, and if a display unit has a line drawing capability, it should be used with this requirement in mind.





## SELECTED BIBLIOGRAPHY

- Adams Associates. Computer Characteristics Quarterly.  
Second Quarter. Watertown, Massachusetts: Keydata  
and Adams Associates, Inc., 1968.
- Brown, Robert G. Statistical Forecasting for Inventory  
Control. New York: McGraw-Hill Book Company, 1959.
- Burdick, Donald S., and Naylor, Thomas H. "Design of  
Computer Simulation Experiments for Industrial  
Systems," Communications of the ACM, 9 (May, 1966),  
329-339.
- Cochran, William G., and Cox, Gertrude M. Experimental  
Designs. New York: John Wiley & Sons, Inc., 1962.
- Hays, William L. Statistics for Psychologists. New York:  
Holt, Rinehart and Winston, 1963.
- International Business Machines Corporation. IBM System/  
360 Operating System, Graphic Programming Services  
for IBM 2250 Display Unit, Form C27-6909-4.  
December, 1967.
- \_\_\_\_\_. IBM System/360 Component Description, IBM 2250  
Display Unit Model 1, Form A27-2701-1. January 27,  
1967.
- Luxenberg, H. R., and Kuehn, Rudolph L., eds. Display  
Systems Engineering. New York: McGraw-Hill Book  
Company, 1968.
- Naylor, Thomas H.; Balintfy, Joseph L.; Burdick, Donald S.;  
and Kong, Chu. Computer Simulation Techniques.  
New York: John Wiley & Sons, Inc., 1966.
- \_\_\_\_\_. ; Werth, Kenneth; and Wonnacott, Thomas H.  
Methods for Analyzing Data from Computer Simulation  
Experiments," Communications of the ACM, 10  
(November, 1967), 703-710.
- Plossl, G. W., and Wight, O. W. Production and Inventory  
Control. Englewood Cliffs, N. J.: Prentice-Hall,  
Inc., 1967.



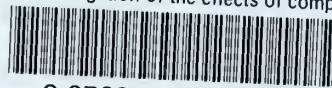
- Sackman, H.; Erikson, W. J.; and Grant, E. E. "Exploratory Experimental Studies Comparing Online and Offline Programming Performance," Communications of the ACM, 11 (January, 1968), 3-11.
- Steel, R. G. D., and Torrie, J. H. Principles and Procedures of Statistics. New York: McGraw-Hill Book Company, 1960.
- Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill Book Company, 1962.





thesP9444

An investigation of the effects of compu



3 2768 001 93036 5

DUDLEY KNOX LIBRARY